ABSTRACT

Title of Thesis: Multiresolution Techniques:
Laplacian Pyramid Coding
and its Comparison with Subband Coding

Danmei Chang, Master of Science in Electrical Engineering, 1991
Department of Electrical and Computer Engineering

Thesis directed by: Dr. Chung H. Lu
Associate Professor
Department of Electrical and Computer Engineering

The multiresolution pyramid and related structures have been developed as intermediate representations between the purely frequency domain and purely spatial domain. In this thesis. we discuss two important pyramid structures which are often used in multiresolution representation: Laplacian pyramid and subband coding.

Laplacian pyramid. a typical pyramid structure. is studied in detail. By means of software implementation. we obtain a comprehensive understanding of its features and applications in image representation and data compression. Several modifications of pyramid structure are used to compress the data dynamic range, reduce the computational complexity and improve performance. Performance of the conventional Laplacian pyramid and the modified pyramid structures are studied and compared.

Multiresolution Techniques:

Laplacian Pyramid Coding and Its Comparison with Subband Coding

by

Danmei Chang

Thesis submitted to the Faculty of the Graduate School of

the New Jersey Institute of Technology in partial fulfillment of

the requirements for the degree of

Master of Science in Electrical Engineering

1991

APPROVAL SHEET

Title of Thesis: Multiresolution Techniques:
Laplacian Pyramid Coding and Its Comparison
with Subband Coding

Candidate: Danmei Chang
Master of Science in Electrical Engineering, 1991

Thesis and Abstract Approved by the Examining Committee:

Dr. Chung H. Lu, Advisor   Date
Associate Professor
Department of Electrical and Computer Engineering

Dr. Edwin Hou   Date
Assistant Professor
Department of Electrical and Computer Engineering

Dr. Yun Q. Shi   Date
Assistant Professor
Department of Electrical and Computer Engineering

New Jersey Institute of Technology, Newark, New Jersey.

VITA

Danmei Chang

434 North Second Street
East Newark. NJ 07029
(201) 484-6206

Date of Birth

Place of Birth

Education

1989-1990	New Jersey Institute of Technology	MSEE
1979-1983	Shanghai Jiao Tong University	BSEE

ACKNOWLEDGMENT

I am very grateful to my graduate advisor Dr. Chung H. Lu whose inspirations and guidances benefited me significantly: without his support this work could not have been finished.

I would also like to thank all of the professors who have taught me, especially Dr. Y. Shi; his 'Digital Image Processing' course introduced me into the wonderful field of image processing.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Recently there has been a lot of interest in the application of human visual models in image processing. Research in the area of vision has shown that there exist mechanisms in biological visual systems that are selective with regard to both spatial and frequency orientation [18]. A good image coding technique should (1) take into account the human visual system model, (2) manipulate the properties of both the spatial and transform domain, (3) achieve high data compression ratio while guaranteeing a faithful reconstruction of the original image.

A variety of image processing techniques have been developed to achieve these general goals in representing and analyzing images over the past 10 years. Among these, multiresolution techniques have become the current trend in image processing and computer vision.

The multiresolution pyramid and related structures have been developed as representations that are intermediate between the purely spatial and purely

frequency domain representations. The basic functions are 'compact' in space and in spatial frequency. Individual pyramid samples represent an image pattern within a neighborhood. rather than at a point (pixel) or over the entire image (frequency component).

Similar multiresolution image representation techniques have been developed in diverse scientific disciplines. Three multiresolution representation approaches have attracted attention for applications in image compression: the Laplacian pyramid, subband decomposition, and most recently. the wavelet transform. These were developed by individuals in different field and with different motivations. yet present day implementations are virtually identical[1].

The Gaussian. or lowpass pyramid is a sequence of reduced resolution copies of the original image. Successive levels of the pyramid are obtained by applying a small kernel lowpass filter to the preceding level. then subsampling the filtered image. The Laplacian. or bandpass pyramid decomposes an image into a set of octave wide bandpass components. Each Laplacian image can be formed as the difference between Gaussian image at same layer and the expanded image of the Gaussian image at the next layer.

Burt and Adelson [1] proposed such a Laplacian pyramid structure which was the first Laplacian pyramid applied to image compression. It provides good compression performance and exact image reconstruction. It has an advan-

tage in that the image analysis can be performed using the same structure as compression [1]. Laplacian pyramid structure [2] plays an important role in the development of multiresolution techniques. This is because pyramid coding combines features of predictive and transform coding methods. its hierarchical structure is similar to that of the nervous system, it uses functions close to those of the human visual system. Image sensed by the eyes are decomposed into bandpass components as they move to the visual cortex of the brain. Independent mechanisms. or channels. within the visual pathway carry visual information in roughly the same way as that in a level of the Laplacian pyramid. Furthermore. information carried by an individual neuron is analogous to a sample in the pyramid [2]. Laplacian pyramid structure in [2] also has the elegant capabilities for progressive transmission or reconstruction. By studying the recent researches in multiresolution image processing. or multi-channel image processing. one would be impressed by the wide applications of similar structures in the decomposition of images. As a matter of fact. Laplacian pyramid coding has significant effect on the later multiresolution techniques. such as subband coding and wavelet transformation used in image compression and pattern recognition.

Subband coding (SBC), originally used for speech coding, has been successfully extended to image coding [31]. The basic idea behind subband coding is to

4

split up the frequency band of the signal and then to code each subband using a coder and bit rate accurately matched to the statistics of that band.

A high data compression rate can be achieved through the subband coding, in addition to this, SBC offers two other advantages. First, the error in coding a subband is confined to that subband, thus exploiting the masking effect of speech. Second, by varying the bit assignment among the subbands, the noise spectrum can be shaped according to the subjective noise perception of the human ear or eye.

Similar to SBC of speech, subband coding of images also suffers from the aliasing problem. Fortunately, it can be cancelled by choosing the suitable filter bank for the ideal case. The design of the filter bank is the most complex part of SBC.

Wavelet representation is another new method for multiresolution signal decomposition. In the wavelet approach to multiresolution decomposition, the difference of the information between the approximation of a function at two different resolutions is computed by decomposing the function into a wavelet orthonormal basis. In practice, we can compute the decomposition of a function into a wavelet orthonormal basis with a quadrature mirror filter (QMF) bank [20].

Multiband image decomposition is also well adapted for coding images because it is possible to match the human visual system sensitivity and take

advantage of the intrinsic statistical properties of images. The sensitivity of human vision depends upon the frequency of stimulus. We want to quantize each frequency band with the minimum number of bits, and at the same time try to reconstruct the best possible image for human visual perception. For this purpose we adapt the quantization noise to the human sensitivities along each frequency band. The more sensitive the human visual system, the less quantization noise is introduced. This enables us to introduce a minimum amount of perceivable distortion in the reconstructed image.

The statistical properties of images give another reason for using multiband decomposition in image coding. It is well known that the intensity of images is locally correlated. Predictive coding has been particularly successful in compressing the number of bits used in coding an image. The wavelet coefficients give a measure of the local contrast at different scale. Since the image intensity is locally correlated, these local contrasts generally have a small amplitude [19], we can take advantage of this property for coding the wavelet coefficients on fewer bits without introducing any noticeable distortion.

The rest of this thesis is organized as follows. Chapter 2 reviews the conventional Laplacian pyramid coding scheme, the implementation of the structure and simulation results are included. The features of the multiscale images are then studied using simulation. This serves as the foundation for the subse-

quent chapters.

Chapter 3 deals with subband coding and wavelet decomposition of images. The brief theoretical analysis of these two approaches is given. then some simulation results are presented for comparing the two methods and to show the similarities and differences between Laplacian pyramid coding and orthogonal decomposition.

Chapter 4 presents some modifications to the conventional image coding. First we introduce a non-linear device — modulo limiter. and its application in predictive system. Then. the possibility of its application in pyramid coding is discussed and a scheme for combining the pyramid structure and modulo limiters is developed.

We then deal with problem of Laplacian coding complexity by providing three modifications to simplify the coding procedures. Finally, we discuss the effect of the interpolation.

Chapter 5 provides the conclusions based on the findings obtained. and suggestions for further research.

Chapter 2

Laplacian Pyramid Coding

The Laplacian pyramid is a versatile data structure that represents an image as a sequence of spatially filtered and decimated versions of the original image. Basically the pyramidal representation involves the construction of a multi-layered structure on which the original image is placed at the bottom, and the successive approximations, each one with a smaller resolution level, are piled on top. Thus, the overall structure resembles a pyramid, every layer having fewer elements than the underlying one and offering a coarser view of the original image [9].

The Laplacian pyramid introduced by Burt and Adelson [2] has the structure shown in Figure 2.1, This scheme combines features of predictive and transform coding. Noncausal prediction of the image is the lowpass filtered copy of the original image. The predicted value for each pixel is computed as a local weighted average. The prediction error is thus the difference between the image and its lowpass filtered copy.
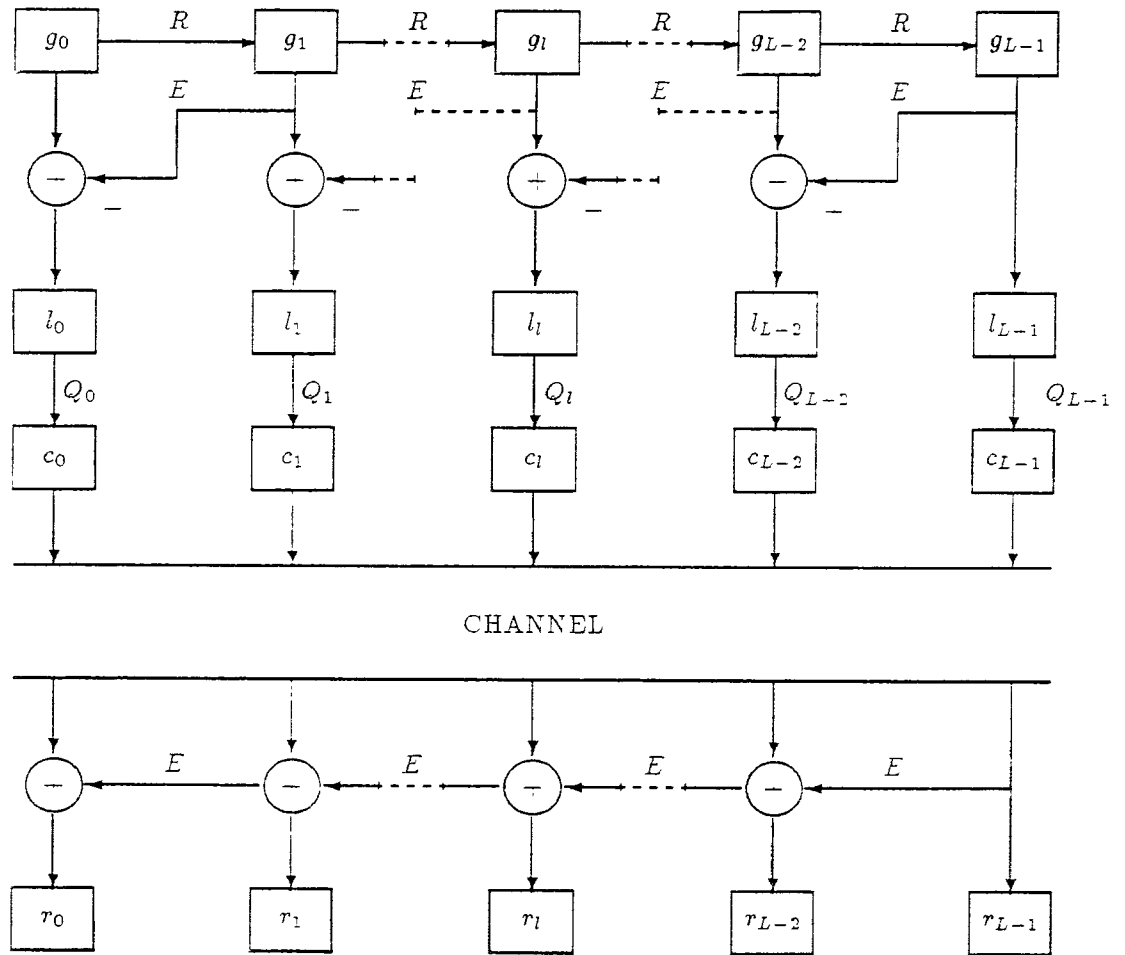
Figure 2.1: Laplacian Pyramid Coding Scheme
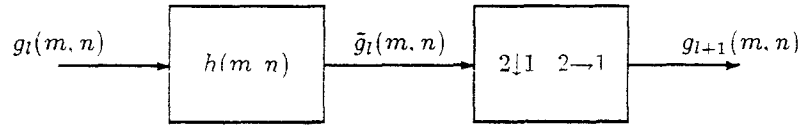
Figure 2.2: Reduction Operation

2.1 Gaussian Pyramid

In the scheme shown in Figure 2.1, the first step is to form a sequence of images g_0, g_1, $\cdots$, g_l, g_{l-1}, $\cdots$, g_{L-1} (L is the number of pyramid layers) by reducing the image in both resolution and sample density. For example, g_{l+1} is obtained by lowpass filtering the image g_l and then subsampling the filtered image. Image g_{l+1} can be considered as a reduced copy of the image g_l, and the procedure to generate g_{l-1} can be called a 'REDUCE' operation. Figure 2.2 shows this reduction operation. where $h(m, n)$ is the impulse response of a two dimensional lowpass filter. and $2\!\downarrow\!1$ denotes the subsampling in the horizontal direction which keeps one column out of two, and $2-1$ is the subsampling in the vertical direction which keeps one row out of two. Usually, the two dimensional filter $h(m,n)$ is a separable symmetric function which can be formed from a one dimensional lowpass filter $\hat{h}(n)$ as

$$h(m,n) = \hat{h}(m)\hat{h}(n) \tag{2.1}$$

In Figure 2.2, the output of the lowpass filter is

$$\tilde{g}_l(m,n) \;\; = \;\; g_l(m,n) \; * \; h(m,n)$$

10

$$= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k,l) g_l(m-k, n-l) \qquad (2.2)$$

If the filter used is an N-tap (N is odd) symmetric lowpass filter, then the above convolution can be written as

$$\tilde{g}_l(m,n) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} h(k,l) g_l(m-k, n-l) \qquad (2.3)$$

After subsampling, the reduced image is thus, equal to

$$g_{l+1}(m,n) = \tilde{g}_l(2m, 2n)$$

$$= \sum_{k=-\frac{N-2}{2}}^{\frac{N-1}{2}} \sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} h(k,l) g_l(2m-k, 2n-l) \qquad (2.4)$$

2.2 Laplacian Pyramid

The Laplacian pyramid is a sequence of error images $l_0, l_1, \cdots, l_{L-1}$. Except on the top, where $l_N = g_N$, the Laplacian image is the difference between two levels of the Gaussian pyramid. Thus, for $0 \leq l < N.2]$

$$l_i = g_l - \text{EXPAND}(g_{l+1}) \qquad (2.5)$$

Here, EXPAND is the reverse of REDUCE (we also use E to denote the EXPAND operator and R to REDUCE operator). It expands the reduced image to the original size by first interpolating new pixel values (usually zeros) between two samples (in both dimensions), and then filters it with the same lowpass filter as used in REDUCE function. Figure 2.3 shows the EXPAND function where 1↑2 puts one column of zeros between two columns, and 1—2 puts one row of zeros between two rows. The image resulting after interpolating zeros can be expressed as

11
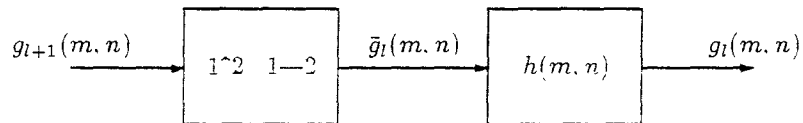
Figure 2.3: Expansion Operation

$$\bar{g}_l(m, n) = \begin{cases} g_{l-1}(\frac{m}{2}, \frac{n}{2}) & \text{if } \frac{m}{2} \text{ and } \frac{n}{2} \text{ are both integers} \\ 0 & \text{otherwise} \end{cases} \qquad (2.6)$$

then. the expanded image is

$$
\begin{aligned}
g_l(m, n) &= \bar{g}_l(m, n) * h(m, n) \\
&= 4 \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h(k, l) \bar{g}_l(m-k, n-l) \\
&= 4 \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{l=-\frac{N-1}{2}}^{\frac{N-1}{2}} h(k, l) g_{l+1}\left(\frac{m-k}{2}, \frac{n-l}{2}\right) \qquad (2.7)
\end{aligned}
$$

The last equation contains only the terms for which $\frac{m-k}{2}$ and $\frac{n-l}{2}$ are both integers. A constant multiplier 4 is applied to correct the DC offset caused by zero insertion.

2.3 Reconstruction

If no quantizer is used, the reconstructed image at top of the pyramid is

$$r_{L-1}(m, n) = g_{L-1}(m, n) \qquad (2.8)$$

For the layer below it, referring to Figure 2.1,

$$r_{L-2}(m, n) = c_{L-2}(m, n) + \text{E}\{r_{L-1}(m, n)\}$$

$$
\begin{aligned}
&= l_{L-2}(m, n) + \mathrm{E}\{g_{L-1}(m, n)\} \\
&= g_{L-2}(m, n) - \mathrm{E}\{g_{L-1}(m, n)\} - \mathrm{E}\{g_{L-1}(m, n)\} \\
&= g_{L-2}(m, n) \qquad\qquad\qquad\qquad\qquad\qquad (2.9)
\end{aligned}
$$

Continuing the procedure, we have

$$
\begin{aligned}
r_l(m, n) &= r_l(m, n) \\
r_{l-1}(m, n) &= g_{l-1}(m, n) \\
\cdots\ \cdots\quad &\quad \cdots\ \cdots \\
r_0(m, n) &= g_0(m, n)
\end{aligned}
$$

The reconstructed image at each level is equal to the corresponding Gaussian images. So, exact reconstruction can obtained as long as there is no channel noise.

2.4   Simulations and Discussions

In order to illustrate the features of the Laplacian pyramid coding, we implement the whole scheme using the computer. By running the programs with different sets of parameters, we have obtained extensive knowledge of this scheme, its advantages and drawbacks, as well as its possible applications and improvements.

First, the Laplacian pyramid as given in [2] is simulated. The two dimensional lowpass filter $h(m, n)$ used here is the same as the one given in [2] which is a separable and symmetric function

$$
h(m, n) = \hat{h}(m)\hat{h}(n) \qquad\qquad\qquad (2.10)
$$

A 5-tap one dimensional filter $\hat{h}(n)$ is used to produce a 5-by-5 lowpass filter.

The filter coefficients are chosen to be

$$\hat{h}(0) \ = \ a = \ 0.4$$

$$\hat{h}(-1) \ = \ \hat{h}(1) \ = \ \frac{1}{4}$$

$$\hat{h}(-2) \ = \ \hat{h}(2) \ = \ \frac{1}{4} - \frac{a}{2} \qquad (2.11)$$

The rest of this chapter includes the observations and discussions about the analysis of the simulated results

Laplacian pyramid and predictive system

Figures 2.4 to 2.7 show the Gaussian images and Laplacian images and their histograms when the number of pyramid levels is 4. Figure 2.4 gives a sequence of Gaussian images. The original image, measures 512 by 512. Each higher level image is half as large in each dimension as the preceeding one. Figure 2.5 illustrates Laplacian images which are taken by subtracting the expanded Gaussian images from the preceeding Gaussian image.

The reduced version of the Gaussian images can be viewed as the predicted images of the original one. and the Laplacian iAmages are then the error images. Actually, a two layer Laplacian pyramid structure is similar to an open-loop DPCM system (D*PCM) [11]. Figure 2.5 shows that the laplacian images have the features that are similar to the error image in a predictive system. In a DPCM system a initial value is the key used to reconstruct the original image from the error image, while in a Laplacian pyramid structure the Gaussian image on the top of pyramid has to be transmitted in order to reconstruct the original image from Laplacian images. Figure 2.6 are the histograms for Gaussian images. and Figure 2.7 are histograms for the first three Laplacian images (the Laplacian image at the top is a Gaussian image). The

14

redundancy is reduced by subtracting a predicted value from each image pixel. The reduction of redundancy results in a concentration of pixel values around zero, and therefore, a reduced variance and entropy. In Laplacian pyramid coding, the noncausal prediction of each pixel based on a neighborhood of the pixel provides a better prediction than the linear, causal prediction. This permits a good data compression rate for code words of. Substantial reduction is realized through quantization (particularly at low pyramid level) and reduced sample density (particularly at high pyramid levels).

A characteristic property of the open-loop DPCM (D^*PCM) is that predicted values at transmitter and receiver (unlike in closed-loop DPCM) are based on different inputs and therefore differ. As a result, the reconstruction error consists of two terms, the quantization noise term plus an additional term equal to the difference between predictor outputs at transmitter and receiver. Quantization error is approximately white noise if the number of quantizer levels is sufficiently high. Hence in D^*PCM, the reconstruction noise is non-white, with each quantization error causing an infinite output sequence. The Laplacian pyramid in Figure 2.1 also suffers from this error propagation problem. This problem can be solved by the quantization noise feedback technique. A Laplacian Apyramid with QNF was proposed in [13]. This scheme has a structure similar to the closed-loop DPCM. An important property of this scheme is that the reconstruction error is identical to the quantizer error.

Coarse-to-fine structure

Laplacian image at the lower layer contains finer detailed information than the ones at the higher layers. These fine details are the features such as edges and bars. In Laplacian pyramid representation, image features are enhanced

Figure 2.4: Gaussian Images
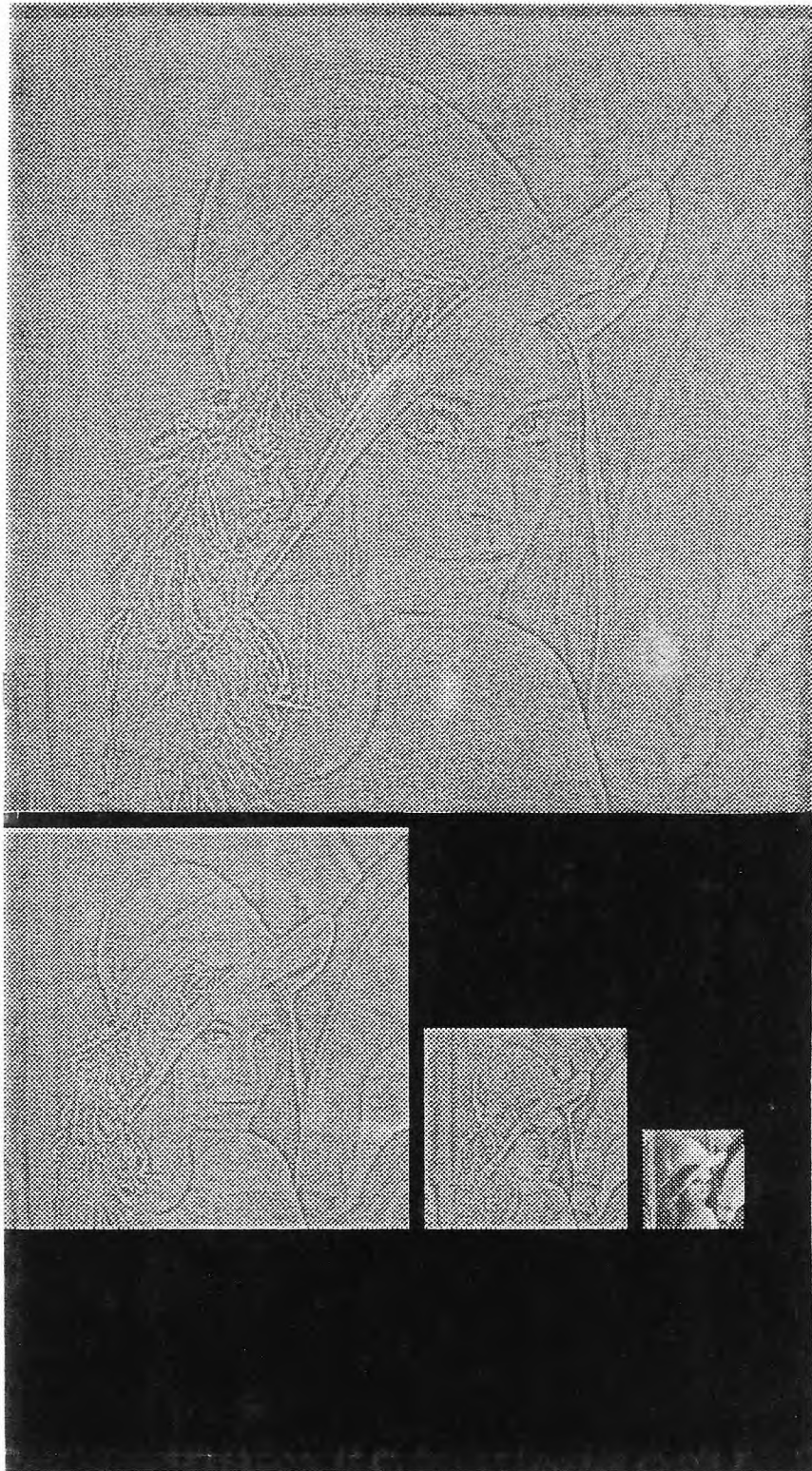
Figure 2.5: Laplacian Images
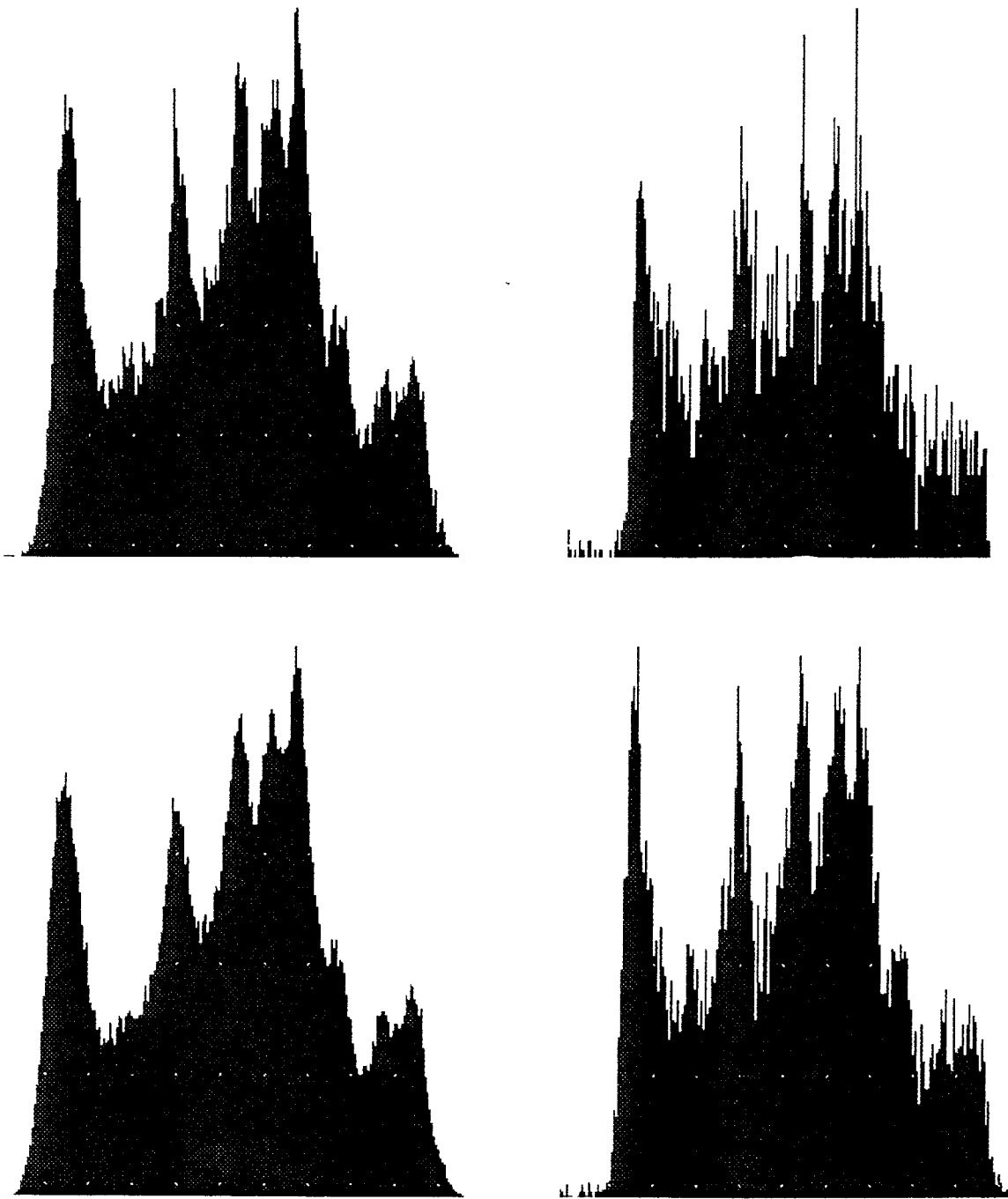
Figure 2.6: Histograms of the Gaussian Images

Figure 2.7: Histograms of the Laplacian Images

and are directly avaliable for various image processing and pattern recognition tasks. Also, because of this coarse-to-fine structure, the Laplacian pyramid is particularly well suited for progressive image transmission by transmitting the coarse version first. The transmission procedure can stop anytime when the reconstruction image from received images satisfied the requirements.

Multichannel representation

Another observation is that when generating the Laplacian pyramid. one automatically has access to quasi-bandpass copies of the image [2] because of the decimation, interpolation and subtraction. Just as the Gaussian images are a set of lowpass filtered copies of the original image, the Laplacian images are a set of bandpass filtered copies of the image. The scale of the Laplacian operator doubles from level to level of the pyramid, while the center frequency of the passband is reduced by an octave [2]. This feature allows the similar pyramid structures to be widely used in the SBC and wavelet transform techniques.

Exact reconstruction

Same structures and filters in encoder and decoder assure exact reconstruction of the original image when there is no channel noise. Lowpass filtering and round-off will not cause reconstruction error. This is a suitable structure for lossless data compression. The symmetric structure makes it easy to design and implement the filter.

Laplacian pyramid and data compression

Table 2.1 lists the entropies for each Gaussian image and Laplacian image, as well as the total bit rate.

Table 2.1: Entropies of the Laplacian pyramid

woman	Gaussian Images	Laplacian Images
level 0	7.445072	4.621148
level 1	7.407004	4.811689
level 2	7.356472	5.396242
level 3	7.280909	7.280909
BPP	9.870367	6.275100

Total BPP for Laplacian images is close to the entropy of the original image. This leads to the conclusion that pure pyramid coding without quantizers will not achieve a high data compression rate. Further data compression will be achieved through proper quantization.

The data compression rate obtained through Laplacian pyramid is approximate to 10:1 by using scale quantizers such as optimum quantizers. The vector quantizer has shown its advantages in pyramid coding [9], but resulting data compression rate is still not satisfactory for some applications. For this reason, Laplacian pyramid coding is classified as a first generation image-coding technique [12].

Quantization and effects of quantization noise

The simulations for Laplacian pyramid coding with quantizers were designed using optimum quantizers. The basic idea for the optimum quantizer appears in [3]. We use a similar method, and the Laplacian image data from ten given images is used so that the data set does not favor certain images. Table 2.2 lists the simulation results including BPP and SNR for quantizers with different numbers of levels and pyramids with different numbers of layers. The SNR between original image and reconstructed image is mainly decided

Table 2.2: BPP and SNR using Optimum Quantizers

Pyramid levels	Quantizer levels	'couple'		'woman'	
		BPP	SNR	BPP	SNR
3	3	1.436	26.294	1.250	29.663
	7	2.556	31.979	1.920	32.364
	11	2.980	33.971	2.631	34.351
	15	3.483	35.398	3.144	35.198
	19	3.557	35.696	3.203	35.434
4	3	1.167	24.453	0.970	26.514
	7	2.026	29.167	1.680	29.036
	11	2.783	31.175	2.422	30.382
	15	3.304	32.057	2.953	30.878
	19	3.390	32.343	3.026	31.026
5	3	1.100	22.575	0.904	24.475
	7	1.971	27.277	1.625	26.726
	11	2.734	28.491	2.374	27.760
	15	3.259	29.018	2.909	28.069
	19	3.347	29.282	2.983	28.170

by the number of pyramid layers and the number of quantization levels. The quality of the reconstructed image also depends on the features of the input image. Table 2.2 shows that the number of pyramid layers is an important parameter. The more pyramid layers. the bigger the BPP and the better is the resulting reconstructed image. Table 2.2 also shows that a 4 layer pyramid is a good compromise between BPP and SNR. The upper image in Figure 2.8 is the reconstructed image for a 4 layer pyramid and 3 level quantization. The reconstruction error is tolerable and the BPP is small. The quantization noise effect on the image in pyramid coding is different from the effect on a reconstructed image in a DPCM system because of the multilevel structure and local non-causal prediction. The lower image in Figure 2.8 is the reconstructed image from DPCM coding with 3 level quantization.

The quantization noise in the lower layer has less effect on the SNR. We can

assign more quantization levels to the Laplacian image on the high layer. In fact, we do not use the quantizer to the top image. This will not increase BPP considerably since number of pixels on the top is much less than number of pixels of the original image. We have show that the quality of the reconstructed image will be reduced greatly if we quantize the image on top of the Laplacian pyramid. A lot of work has been done on reducing the quantization noise in pyramid coding [28] [13]. By quantization feedback techniques, the quantization noise can be reduced more or less. However, it is difficult to achieve optimum quantization with quantization noise feedback.

Computational complexity

One drawback of the pyramid coding is that the total number of the samples for Laplacian images are more then the samples of the original image. If original image measures N-by-N and the number of pyramid levels is L. then the total number of samples for transmitting is

$$n_s = \sum_{l=0}^{L-1} \frac{1}{2^{2l}} N^2 \qquad (2.12)$$

This number is close to $\frac{4}{3} N^2$ when L is large.

Laplacian pyramid coding requires large amount of additions and multiplications. due to its multilevel structure, and because of the decimation and interpolation operations. The other problem with Laplacian pyramid coding is that the dynamic range of the Laplacian images is twice that of the original image.

The coarse-to-fine feature and flexible structure makes pyramid coding a suitable method for image representation. Many researchers have contributed

Figure 2.8: Reconstructed Images from LP and DPCM

24

their modifications to this coding scheme. After careful study of pyramid coding and combining some other techniques used in signal processing or image processing, Chapter 4 will include some modifications we have worked out. Some further discussion on the performance and applications of the pyramid coding will be included in chapter 4.

Chapter 3

Subband Coding of Images

3.1 Subband Coding of Images

Subband coding is an encoding technique to divide the input signal into a number of separate frequency components. and to encode each of these components separately. This division into frequency components reduces the redundancy in the input and provides a set of uncorrelated inputs to the channel. The *frequency domain* coding techniques have the advantage that the encoding accuracy is always placed where it is needed in the frequency domain. In fact. bands with little or no energy may not be encoded at all [11].

Figure 3.1 illustrates a basic block diagram of the subband coder for with one layer. In the subband coder, the input signal is divided into typically four or more subbands by a modulation process equivalent to single-side-band amplitude modulation. It is then sampled (or subsampled) and encoded. In this process, each sub-band can be encoded according to perceptual criteria that are specific to that band. In receiver, the subband signals are decoded and modulated back to their original locations. They are then summed to give a close replica of the original signal.

Encoding in subbands offers several advantages. By appropriately allocating
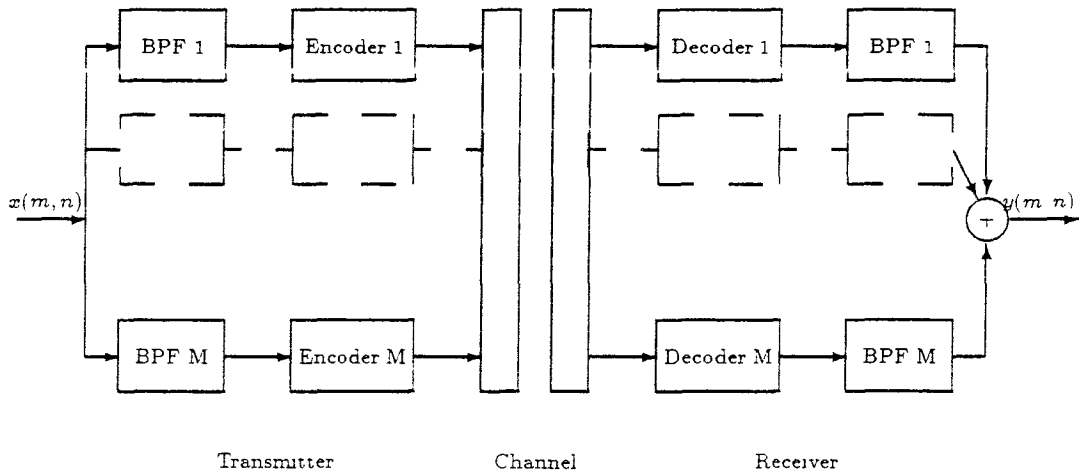
Figure 3.1: Subband Coding

the bits in different bands. the number of quantizer levels and hence reconstruction error variance can be separately controlled in each band. and the shape of the overall reconstruction error spectrum can be controlled as a function of frequency. Take speech coding for example. In the lower frequency bands. where pitch and formant structure must be accurately preserved, a large number of bits/sample can be used: whereas in upper frequency bands, where fricative and noise-like signals occur, fewer bits/sample can be used. Further. quantization noise can be contained within bands to prevent masking of a low-level input in one frequency range by quantizing noise in another frequency range[11].

The most complex part of the coder is the filter bank [11]. The overlapping subband suggests that aliasing effects can occur. This problem is very elegantly tackled in the *quadrature—mirror filter* (QMF) bank approach [11]. In order to cancel aliasing effects when reconstructing the subband signals, the

27

QMF filter bank should have the following properties:

- lower band filter $h_l(n)$ and upper band $h_u(n)$ are respectively symmetrical and anti-symmetrical, FIR designs have an even number of taps. i.e.,

$$h_l(n) = h_u(n) = 0 \qquad n < 0 \ and \ n \geq N \qquad (3.1)$$

$$h_l(n) = h_l(N - 1 - n) \qquad n = 0.1. \cdots, N/2 - 1 \qquad (3.2)$$

$$h_u(n) = -h_u(N - 1 - n) \quad n = 0.1, \cdots, N/2 - 1 \qquad (3.3)$$

- QMF filter bank should satisfy the condition

$$h_u(n) = (-1)^n h_l(n), \quad n = 0, 1. \cdots . N - 1 \qquad (3.4)$$

- amplitude spectrum satisfy the condition

$$|H_l(e^{j\omega})|^2 + |H_u(e^{j\omega})|^2 = 1 \qquad (3.5)$$

Subband coding of images is similar to SBC of speech. The most important step is to design the 2-D filter banks. Wood and O'Neil [31] introduced an approach in which a 2-D QMF bank and pyramid structure are used to decompose the input image. 1-D FIR QMF is extended to 2-D case to approximate the ideal subband filter. Separable filters are sufficient (not necessary) for the most natural four-band extension of the standard two-band QMF filters. A 4-band splitting subband coding system is shown in Figure 3.2 [31]. Where the two dimensional filters $h_{ll}(m, n)$ to $h_{uu}(m, n)$ are obtained from 1-D filters $h_l(n)$ and $h_u(n)$ as

$$h_{ll}(m, n) = h_l(m) h_l(n) \qquad (3.6)$$

28
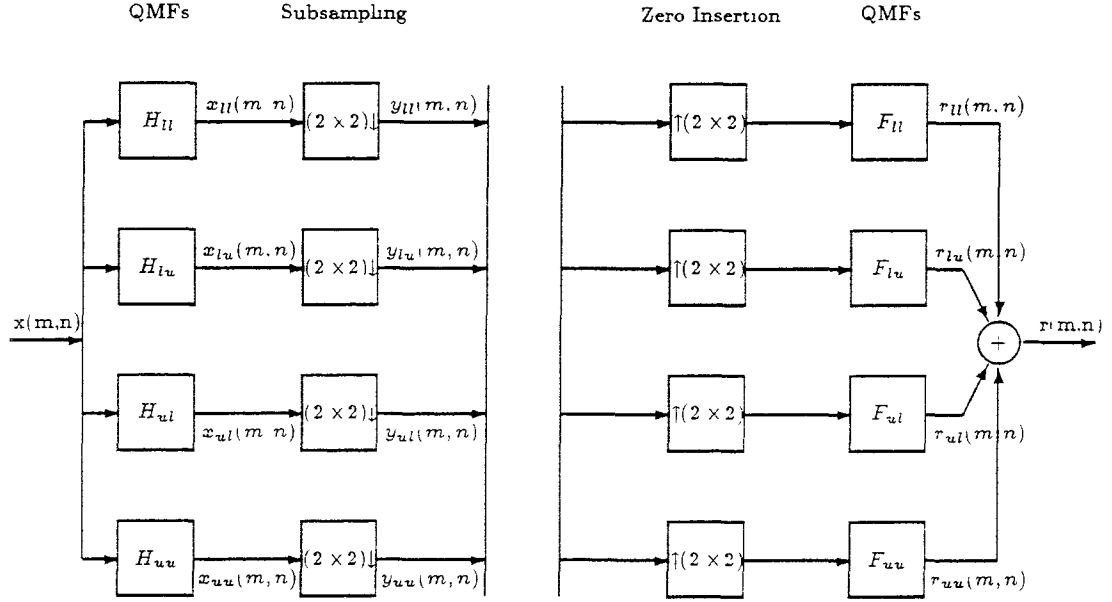
Figure 3.2:   4-Band Subband Coding of Image

$$h_{lu}(m, n) = h_l(m)h_u(n) \tag{3.7}$$

$$h_{ul}(m, n) = h_u(m)h_l(n) \tag{3.8}$$

$$h_{uu}(m, n) = h_u(m)h_u(n) \tag{3.9}$$

and. for real h_{ij}, we have

$$H_{lu}(\omega_v, \omega_h) = H_{ll}(\omega_v, \omega_h + \pi) \tag{3.10}$$

$$H_{ul}(\omega_v, \omega_h) = H_{ll}(\omega_v + \pi, \omega_h) \tag{3.11}$$

$$H_{uu}(\omega_v, \omega_h) = H_{ll}(\omega_v + \pi, \omega_h + \pi) \tag{3.12}$$

From Figure 3.2, we have

$$Y_{ij}(\omega_v, \omega_h) = \frac{1}{4}\sum_{k=0}^{1}\sum_{l=0}^{1} H_{ij}\left(\frac{\omega_v + k\pi}{2}, \frac{\omega_h + l\pi}{2}\right) X\left(\frac{\omega_v + 2k\pi}{2}, \frac{\omega_h + 2l\pi}{2}\right). \tag{3.13}$$

29

and,

$$R_{ij}(\omega_v, \omega_h) = Y_{ij}(2\omega_v, 2\omega_h) F_{ij}(\omega_v, \omega_h) \tag{3.14}$$

The final output is then

$$R(\omega_v, \omega_h) = \frac{1}{4} \sum_{a,b} \left(X(\omega_v + a\pi, \omega_h + b\pi) \sum_{i,j} [H_{ij}(\omega_v + a\pi, \omega_h + b\pi) F_{ij}(\omega_v, \omega_h)] \right) \tag{3.15}$$

It has been shown [31] that the aliasing will be cancelled when the filter bank in the receiver satisfies the following condition.

$$F_{ll}(\omega_v, \omega_h) = 4H_{ll}(\omega_v, \omega_h) \tag{3.16}$$

$$F_{lu}(\omega_v, \omega_h) = -4H_{lu}(\omega_v, \omega_h) \tag{3.17}$$

$$F_{ul}(\omega_v, \omega_h) = -4H_{ul}(\omega_v, \omega_h) \tag{3.18}$$

$$F_{uu}(\omega_v, \omega_h) = 4H_{uu}(\omega_v, \omega_h) \tag{3.19}$$

The simulation of this scheme gives the results in Figure 3.3 to 3.5. When implement the subband coding of images, we use the 1-D filter in horizontal and vertical directions respectively. The 16-band splitting is implemented by doing 4-band splitting to each subband image again.

Figure 3.3 is the original image. Figure 3.4 shows the 4-band subband images, and Figure 3.5 shows the 16-band subband images. Figure 3.4 and 3.5 show that the subband image has a certain orientation selectivity. In Figure 3.4, the upper left image is the lowpass filtered coarse image, the upper right is the image filtered by $h_{lu}(m, n)$, which contains the horizontal detail information, the lower left is the image filtered by filter $h_{ul}(m, n)$ which contains the vertical detail information, and the lower right is the image with highest frequency which shows the diagonal information.

Figure 3.6 shows Laplacian images. Note that Laplacian images do not have orientation selectivity.
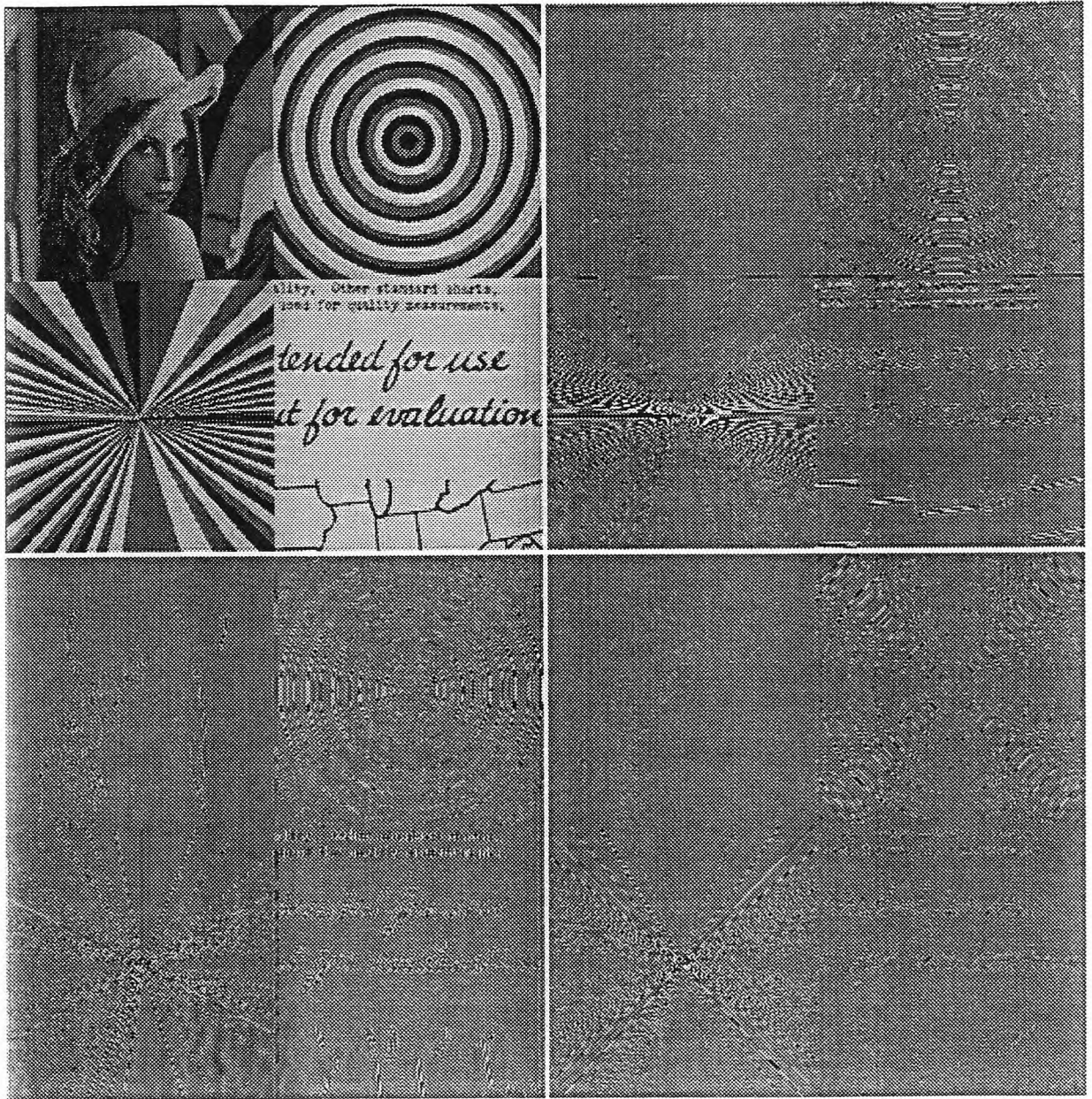
Figure 3.3: Original Image

Figure 3.4: 4-Band Subband Images
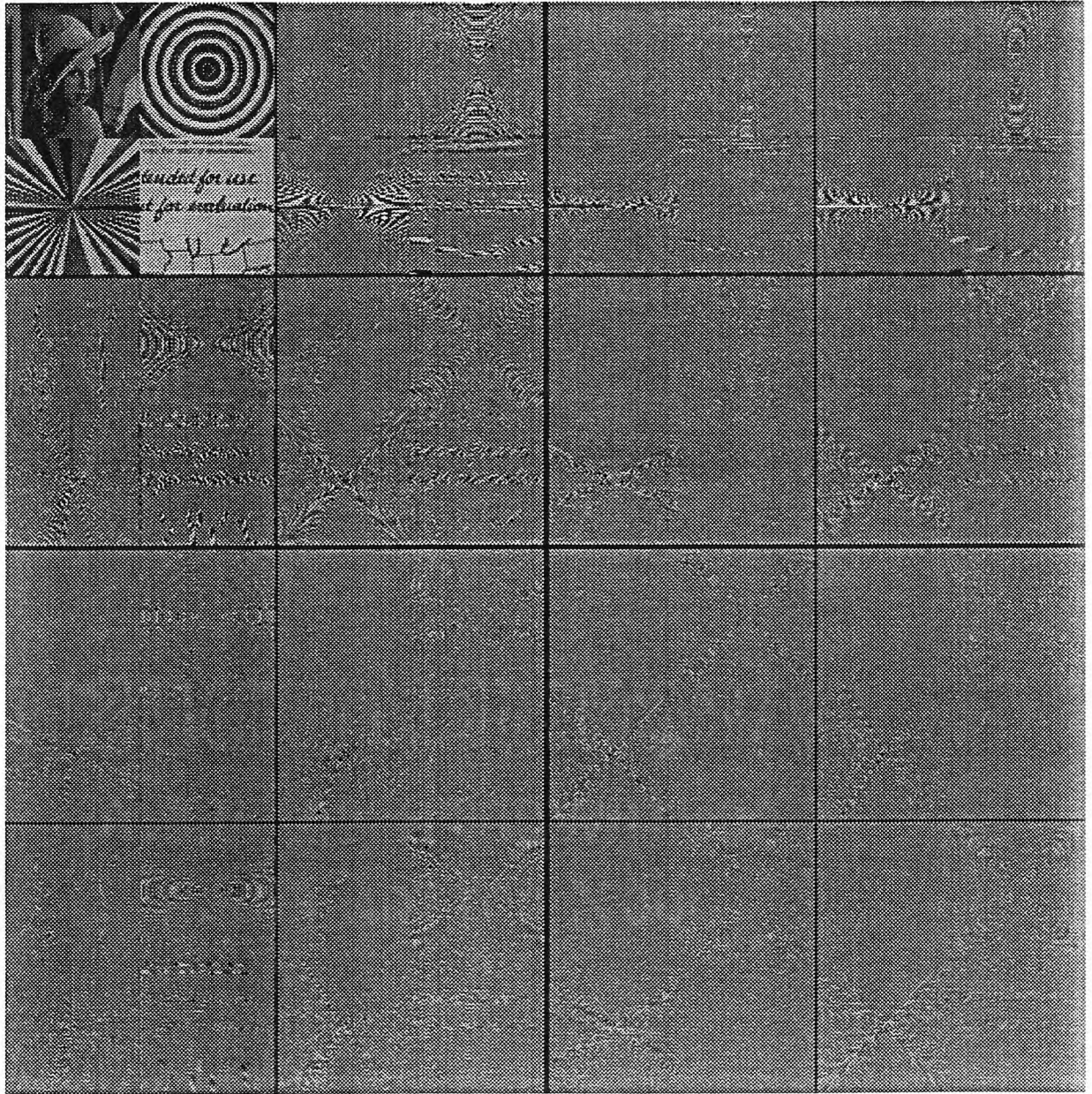
Figure 3.5: 16-Band Subband Images

33
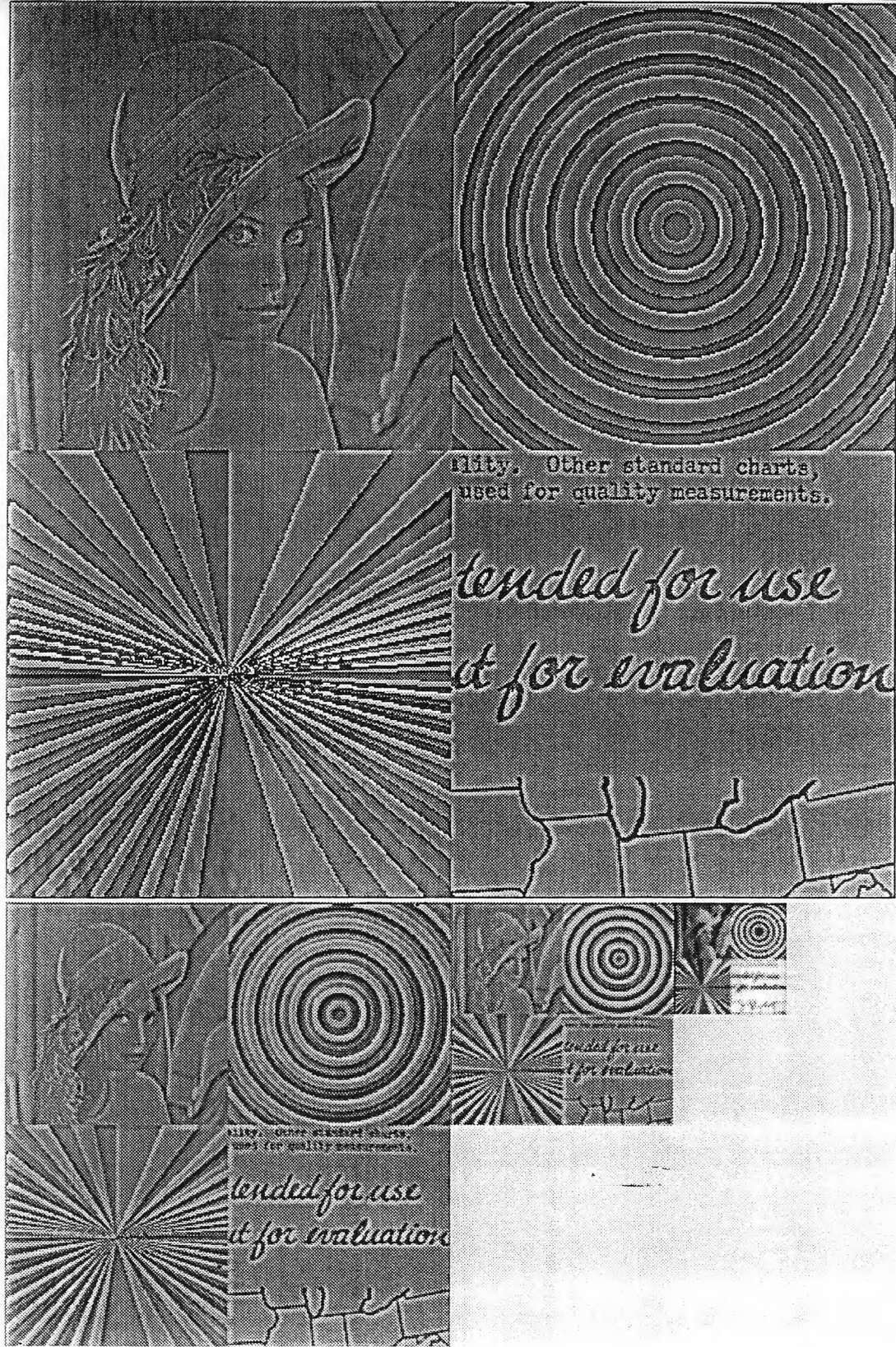
Figure 3.6: Laplacian Images

3.2 Wavelet Decomposition of Images

Wavelet decomposition of an image mean decomposing the image into a family of images on a wavelet orthonormal basis. Morlet [7] defined the wavelet transform by decomposing the signal into a family of functions which are the translation and dilation of a unique function $\psi(x)$. The function $\psi(x)$ is called a wavelet and corresponding wavelet family is given by $(\sqrt{s}\psi(s(x-u)))_{(s\ u)\in\mathbf{R}^2}$. The wavelet transform of a function $f(x) \in \mathbf{L}^2(\mathbf{R})$ is defined by

$$W f(s,u) = \int_{-\infty}^{+\infty} f(x)\sqrt{s}\psi(s(x-u))dx \qquad (3.20)$$

As we mentioned before, with pyramid coding of an image the number of samples is increased by a factor of $\frac{1}{3}$. The multiresolution wavelet model decomposes images into multiresolution components while maintaining the same number of pixels as the original images, and the exact difference can be extracted by decomposing the image into a wavelet orthonormal basis. Mallat [20] proved that in order to do so, the filter should satisfy the condition

$$|H(e^{-J\omega})|^2 + |H(-e^{-J\omega})|^2 = 1 \qquad (3.21)$$

The discrete filters whose transfer function satisfy the above condition are QMF.

At present, the implementation of multiresolution wavelet model using QMF is the same as the subband coding of images.

3.3 Comparisons between Laplacian pyramid coding and Subband Coding

Laplacian pyramid, subband coding and wavelet decomposition are three important multiresolution decomposition approaches. They have some features
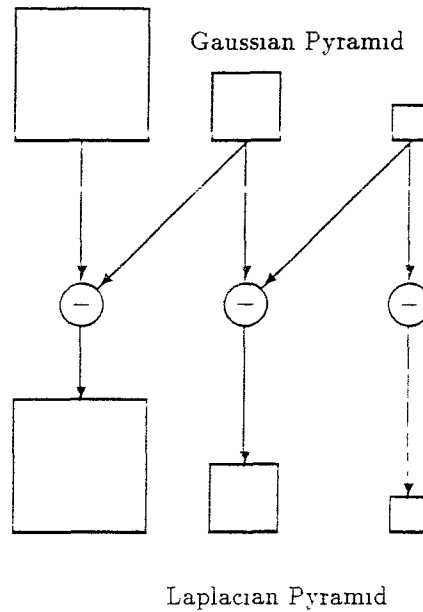
Figure 3.7: Laplacian Pyramid Structure

in common. First, they all decompose the input images into multichannels, and each channel has the bandpass characteristic. By appropriately assigning bits in different channels, the number of quantizer levels and hence reconstruction error can be separately controlled in each of the channels according to the features of the decomposed signals. The high data compression performance can be obtained through this method.

There are three different structures which are often used to represent images with multiresolution. Figure 3.7 is the structure which is used in the Laplacian pyramid representation of the images. And Figure 3.8 shows these two trees structures which are mostly used in subband coding or wavelet decomposition. Structure (a) is a symmetric tree structure, while (b) uses an asymmetric tree structure. If the majority of the image power is contained in lower frequency band, the structure (b) is reasonable and will reduce the coding complexity.
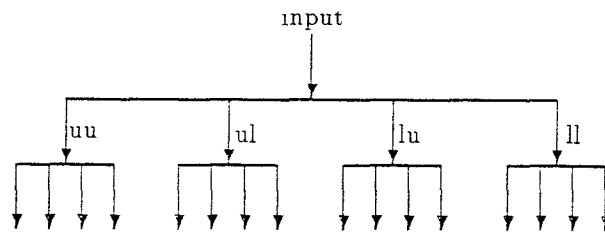
(a)
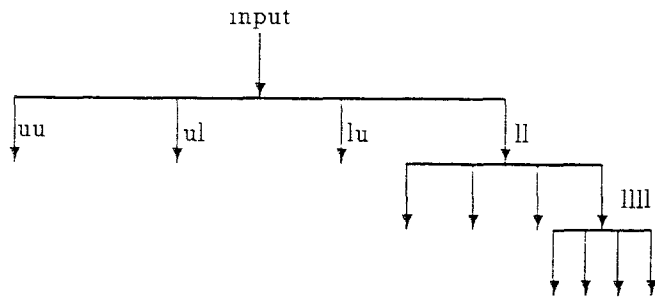
(b)

Figure 3.8:   Tree Structures of Image Decomposition

The decomposition with these tree structures will remain the same number of pixels as the original images.

The Multiresolution concept and QMF are directly related to wavelet orthonormal bases. The subband coding scheme in Figure 3.2 can be considered as a special case of wavelet transform.

Figure 3.9 gives histgrams for the 4-band subband images. The orthogonal decomposition by QMF generates a sequence of images in which the redundancy is removed efficiently. This feature permits subband coding and wavelet decomposition to achieve a very higher data compression rate.

In Laplacian pyramid coding, the encoder and decoder structures are identical, which makes it easy to design filters. We have shown that a 5-tap filter is adequate in our case. The filter bank used in subband coding is much more complex. The filter bank is difficult to design and implement, the filtering procedures require more computations.

The identical structure in transmitter and receiver of a Laplacian pyramid also permits the exact reconstruction of the original image. But for subband coding, the exact reconstruction can be achieve only when the infinity word length is used.

QMF decompose the image into different channels each with certain orientation selectivity, while the Laplacian images combine the edge information in all directions.

In Laplacian coding, the coarse image on top of the pyramid is usually assigned to a greater bit rate. This will not increase the total bit rate too much since the size of this image is much smaller than the size of all other Laplacian images. The symmetric tree structure subband coding will generate a lowpass filtered image with the same size as all of the other subband images. The bit

rate of lowpass images will have the same effect as the other images have.
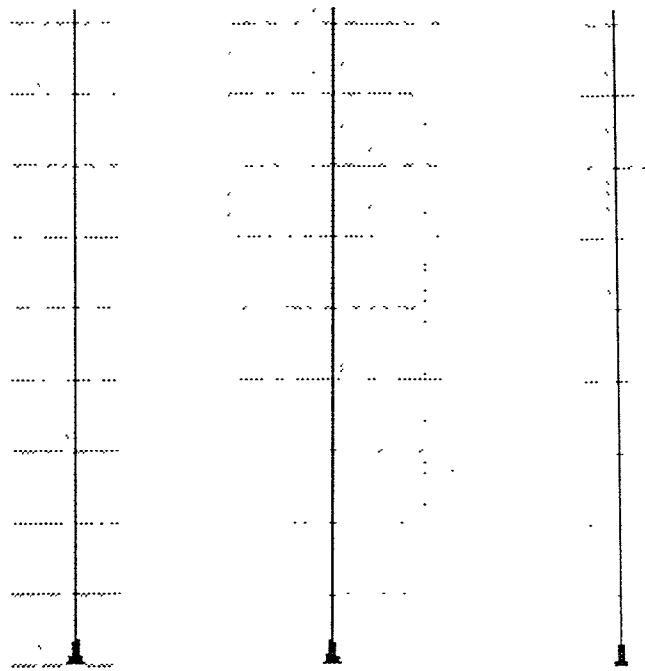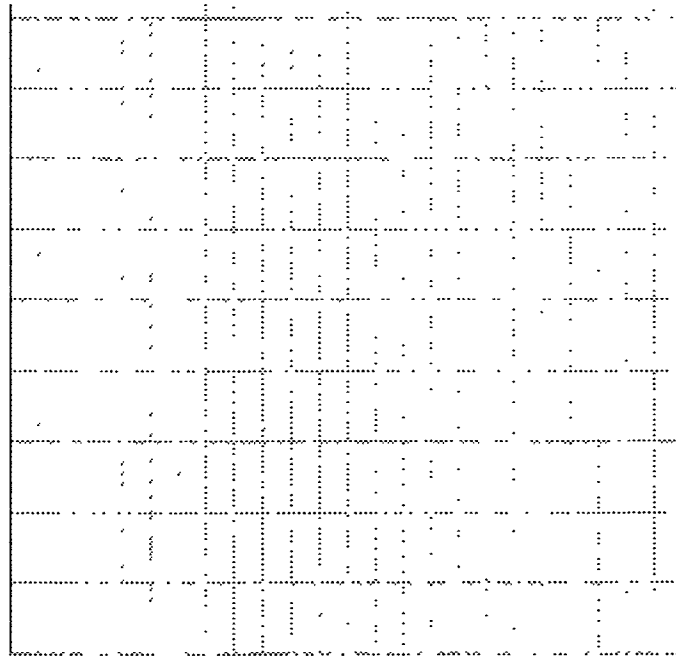
Figure 3.8: Histgrams of 4-band subband images

40

Chapter 4

Modifications to Laplacian Pyramid Coding

In this chapter we will propose some modifications to the pyramid coding. and compare the features among the different structures.

4.1 Compressing Dynamic Range of the Laplacian Images

In Laplacian pyramid coding, the Laplacian images are obtained by subtracting the two Gaussian images. So the gray level range for Laplacian image is twice as much as the original input image. Suppose the input image is a 8-bit gray level image, the Laplacian image will need 9 bits for directly storing or transmitting each pixel when perfect reconstruction is required. Unless otherwise stated, all the images used as input images are digitized with a 512 by 512 raster and quantized to 256 levels, thereby permitting an 8-bit representation. In the following section. we will deal with a method to compress the dynamic range of the Laplacian images. We will see that a modulo limiter can solve the problem efficiently.
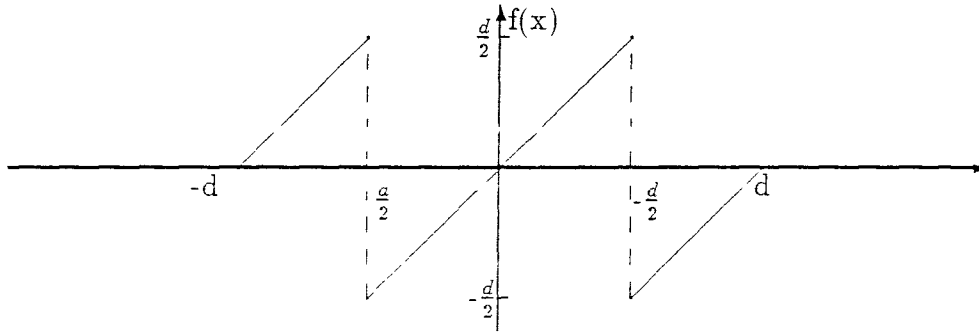
Figure 4.1: Modulo Limiter

4.1.1   Introduction to Modulo Limiter

A modulo limiter (ML) is a nonlinear device which repeatedly applies a preset offset to the input until the output falls within a preset range [15]. The relation between the input and output is given in Figure 4.1. Mathematically, the modulo limiter can be represented in terms of a modulo function. Its digital implementation is very simple. In actuality, application of the modulo limiter in a digital system results in the need for less hardware [15].

Modulo limiting has been successfully applied in PCM system for speech coding (MPCM) [27,22]. The modulo limiter has also been applied in DPCM system for image coding [15].

In predictive coding systems, the error signal is obtained by subtracting the predicted value from the original signal. If the dynamic range of the original signal is d, dynamic range of the error signal will then be doubled. The error signal usually remains about zero if the predictor is well designed, in this case one can simply truncate the error signal to limit it to the same dynamic range as the original signal and the output will not have significant distortion. This limitation is particularly useful for image processing, since the input images we use are mostly character images in which each pixel value is stored as a

42
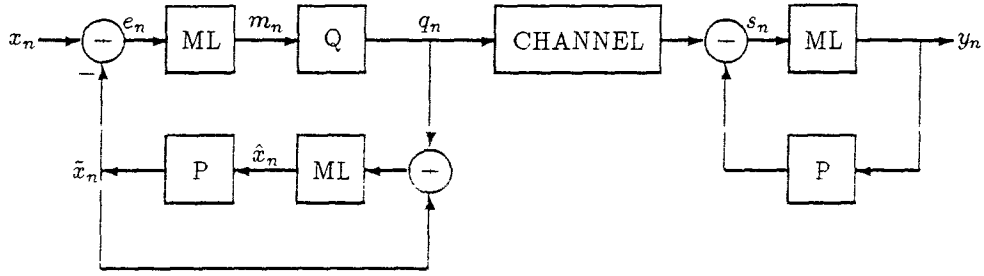
Figure 4.2: DPCM with Modulo Limiting

byte. By using modulo limiters. one can reduce the dynamic range of the error image without introducing error.

Figure 4.2 shows a DPCM system with modulo limiters. One modulo limiter at the transmitter compresses the dynamic range of the prediction error to ease the design of the quantizer and consequently to improve coding efficiency [15]. Another limiter at the receiver expands the compressed prediction error to the original range. Modulo limiters can also be used to improve dynamic range or reduce noise level of a signal [4].

The modulo limiter can be represented by the relation

$$f(x) = (x - \frac{d}{2})\text{MOD}\, d - \frac{d}{2} \tag{4.1}$$

Let x_n be the ith input sample and $\hat{x}_n$ be the nth reconstructed sample. The prediction $\tilde{x}_n$ is a function of the reconstructed value $\hat{x}_{n-1}, \hat{x}_{n-2}, \hat{x}_{n-3}, \cdots$, that is

$$\tilde{x}_n = P(\hat{x}_{n-1}, \hat{x}_{n-2}, \hat{x}_{n-3}, \cdots) \tag{4.2}$$

The prediction error is therefore.

$$e_n = x_n - \tilde{x}_n \tag{4.3}$$

If both the input and reconstructed values are limited in the range $[-\frac{d}{2}, \frac{d}{2})$, the range of the prediction error e_n will be $[-d, d)$. Using equation (4.1), output

43

of the modulo limiter is

$$
\begin{aligned}
m_n &= f(e_n) \\
&= (e_n + \frac{d}{2})\mathrm{MOD}\, d - \frac{d}{2} \\
&= e_n + k\, d \qquad\qquad\qquad\qquad\qquad (4.4)
\end{aligned}
$$

where

$$
k = \begin{cases}
1 & -d \le e_n < -\frac{d}{2} \\[2mm]
0 & -\frac{d}{2} \le e_n < \frac{d}{2} \\[2mm]
-1 & \frac{d}{2} \le e_n < d
\end{cases}
$$

Assume the quantization noise for sample e_n is n_n, then the reconstructed value $\hat{x}_n$ is

$$
\begin{aligned}
\hat{x}_n &= f(e_n + k\, d + n_n + \tilde{x}_n) \\
&= f(x_n - k\, d + n_n) \\
&= f(x_n + n_n) \qquad\qquad\qquad\qquad (4.5)
\end{aligned}
$$

If n_n is negligible. the above representation for $\hat{x}_n$ can be written as

$$
\hat{x}_n = x_n \qquad\qquad\qquad\qquad (4.6)
$$

Therefore, in the case of no qnantization noise the reconstructed value $\hat{x}_n$ is equal to the input value x_n. This means that a DC offset which may be added by the first ML. will be automatically corrected by the second ML. The prediction-error modulo limiter reduces the dynamic range of the prediction error by 6 dB.

One problem with the modulo limiter is that when quantization noise exists and the value of e_n is around the break points, the DC offset might not be corrected properly.

44

To use a modular limiter in a predictive system, first we need a feedback circuits, and second we have to use an identical structure with that of the feedback mechanism in the decoder. Figure 4.2 is an example that the output of the decoder y_n is equal to the input signal x_n when channel is noise free and quantization noise can be neglected.

4.1.2   Modulo Limiters in Pyramid Coding

As we mentioned before, the Laplacian pyramid possesses features of the predictive system. The prediction value is a local weighted average instead of the combination of the previous pixels as in the DPCM. Besides the features such as multiresolution and realtime processing, pyramid coding is a lot like DPCM system. The Laplacian images are similar to the error image in DPCM, and its dynamic range is twice that of the input image.

These features allow us to use the modulo limiters in a Laplacian pyramid. Following the analysis for the DPCM with the modulo limiter, we construct the following scheme which combines the Laplacian pyramid with modulo limiters. The Figure 4.3 shows the application of ML in Laplacian pyramid, for simplicity we only give the first four levels here.

Let's consider a pyramid structure as shown in the Figure 4.3 first. The result can be easily extended to the pyramid with arbitrary number of levels. In the following proof, g_l denotes the Gaussian image in pyramid level l and $g_l(m, n)$ is the gray level value for the pixel at position (m, n). Similar notations apply for Laplacian image l_l, the quantized Laplacian image c_l and the reconstructed image r_l. The original image is stored as an 8-bit character image, that is $g_0(m, n) \in [0, 256)$, since the REDUCE function keeps the dynamic range of the images. The error image $g_l(m, n)$ - EXPAND$\{g_{l+1}(m, n)\} \in [-256, 256)$,
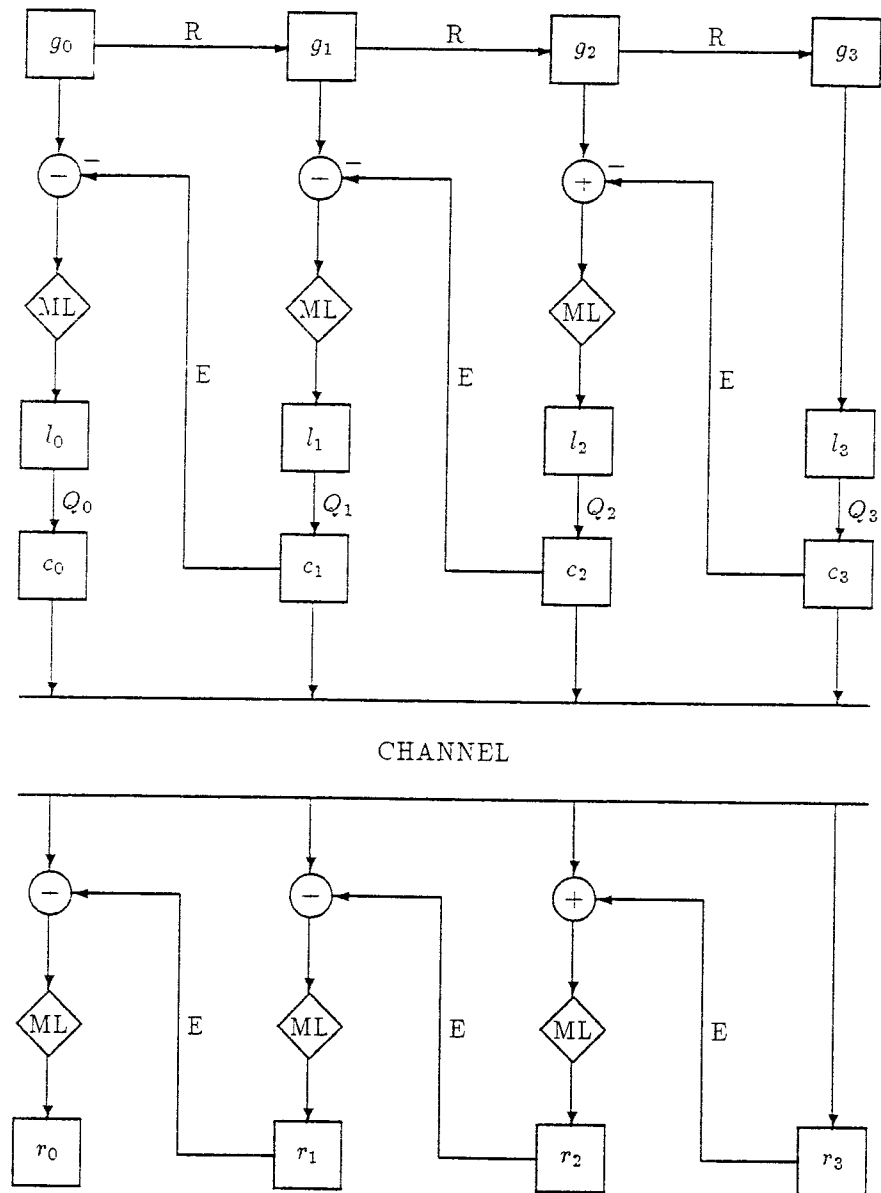
Figure 4.3: Pyramid with Modulo Limiter

where EXPAND$\{g_{l+1}(m,n)\}$ is the result from the EXPAND of $g_{l+1}(m,n)$.
We are now ready to prove that the modulo limiter can be used in a pyramid
coding scheme by first assuming no quantizer or negligible quantization noise.
On the top level we have

$$r_3(m,n) = c_3(m,n) = g_3(m,n) \tag{4.7}$$

Then, for the level below it, we have

$$
\begin{aligned}
c_2(m,n) &= f(g_2(m,n) - \mathrm{E}\{g_3(m,n)\}) \\
&= g_2(m,n) - \mathrm{E}\{g_3(m,n)\} - k \cdot 256
\end{aligned}
\tag{4.8}
$$

where
$$
k = \begin{cases}
-1 & 128 \le g_2(m,n) - \mathrm{E}\{g_3(m,n)\} < 256 \\
0 & -128 \le g_2(m,n) - \mathrm{E}\{g_3(m,n)\} < 128 \\
1 & -256 \le g_2(m,n) - \mathrm{E}\{g_3(m,n)\} < -128
\end{cases}
$$

In the decoder,

$$
\begin{aligned}
r_2(m,n) &= f(c_2(m,n) + \mathrm{E}\{r_3(m,n)\}) \\
&= f(g_2(m,n) - \mathrm{E}\{g_3(m,n)\} + k.256) + \mathrm{E}\{r_3(m,n)\} \\
&= g_2(m,n)
\end{aligned}
\tag{4.9}
$$

This shows that in a Laplacian pyramid with modulo limiters, the ML in the
decoder can also automatically correct the DC offset which may be introduced
by the ML in the encoder.

By using the same procedure, we can prove at last

$$r_0(m,n) = g_0(m,n) \tag{4.10}$$

Although we derived the above conclusion from a 4-level pyramid. This con-
clusion remains true, if and only if the following condition are satisfied

- The effect of the quantizers is negligible. On the top of the pyramid we have $r_{L-1}(m,n) = g_{L-1}(m.n)$

- REDUCE and EXPAND operations do not change dynamic range of the images.

- Modulo limiters should appear in pairs. one in encoder and other in decoder.

- Channel noise can be ignored.

The computer simulations also confirm the conclusion we obtained here. It is very easy to implement the modulo limiter using a computer program. Actually we are doing the modulo computation when we take lower 8 bits from a 9-bit number. When quantizers are used the modular limiter reduces the input's dynamic range and makes the design of the quantizer easier because it needs fewer levels. The problem arises when there are very few quantization levels. in this case the quantization noise will cause an error in correction. This occurs when the ML in the decoder doesn't compensate for the DC offset.

Even with these problems, the modulo limiter used in the Laplacian pyramid is particularly useful for applications such as lossless progressive transmission, or lossless variable word length coding. The modulo limiters reduce the dynamic range of the signal to be transmitted or quantized, thus it improves the data compression gain. Improvement is substantial for some images, but can be minimal for others: in the other words, image signals are non-stationary. Obviously, the ML will help when the prediction error is large. The noncausal prediction based on a symmetric neighborhood centered at each pixel, yields a more accurate prediction. hence the compression rate by modulo limiter will usually be less then the rate in DPCM_ML.

$$g_{l+1}(m, n)$$

```
┌─────────────┐        ┌─────────────┐
│  2↓1   2→1  │        │  2↓1   2→1  │
└─────────────┘        └─────────────┘

┌─────────────┐        ┌─────────────┐
│   h(m,n)    │        │   h(m,n)    │
└─────────────┘        └─────────────┘
                                  −
g_l(m, n) ───────────────────────(−)──────→  l_l(m, n)
```
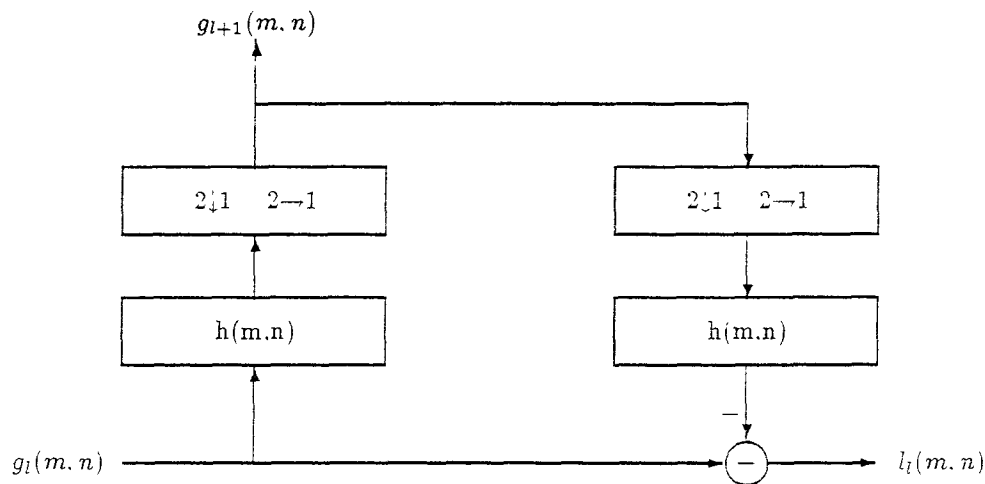
Figure 4.4: Laplacian Model

It is not necessary to use modulo limiter pairs at every layer of the pyramid, we need to can use limiters only in the layers in which the pixel gray values do not concentrate around zero. On the top of pyramid the application of modulo limiters will not cause any difference.

4.2 Reducing Computational Complexity

Although filtering in the Laplacian pyramid can be performed efficiently with a fast algorithm, it still needs too much computation compared to the predictive systems such as a DPCM system. In the conventional Laplacian pyramid data structure, two lowpass filtering operation are required to generate a Laplacian image. Figure 4.4 displays the procedures for generating a Laplacian pyramid at level l. We call it the Laplacian model.

We will discuss the possibility of reducing the computation complexity in pyramid coding. Some schemes with simpler filtering approaches are developed

49
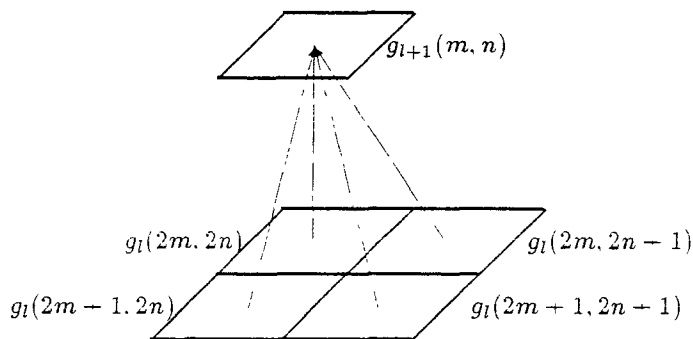
Figure 4.5: Modification 1

and the features are compared with the original scheme.

4.2.1 Modification 1—Mean Filtering

A mean pyramid is formed by successively averaging over 2×2 neighboring nodes as shown in Figure 4.5. Here the averaging is a lowpass filtering of the image.

At pyramid level $l - 1$. each pixel is formed by

$$g_{l+1}(m,n) = \frac{g_l(2m,2n) - g_l(2m,2n+1) + g_l(2m+1,2n) + g_l(2m+1,2n+1)}{4}$$

(4.11)

Comparing with Figure 4.4. we avoid the two convolution operations for generating the error images. Computer simulation shows the mean pyramid data structure has features similar to conventional pyramid structure. For certain images this scheme can achieve features which are even better than the conventional pyramid scheme.
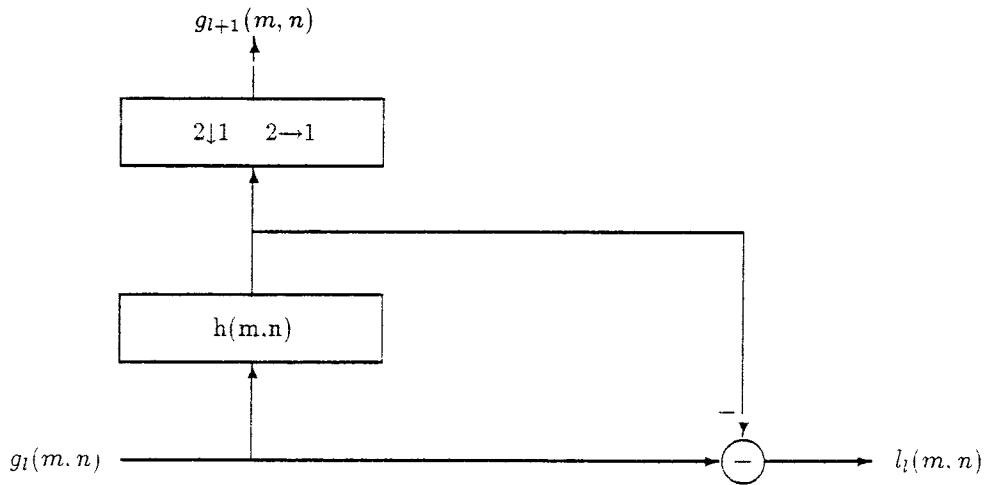
Figure 4.6: Modification 2

4.2.2 Modification 2

By examining the procedures for generating Laplacian images in Figure 4.4. we find the image l_l is obtained by subtracting the lowpass filtered copy of the Gaussian image from g_l. We can avoid the EXPAND operation by using the scheme as illustrated in Figure 4.6. The problem with this procedure is that we can not obtain the exact reconstruction because the input of EXPAND in the decoder is different from the one in encoder, but the distortion is very small and is hardly noticeable.

4.2.3 Modification 3

Another scheme shown in the Figure 4.7 makes use of the correlation feature in the image when generating a Gaussian image. The Gaussian image is taken by only decimating the image below it. The computation complexity of this scheme is similar to modification 2.

$g_{l+1}(m, n)$

| $2{\downarrow}1$ $2{\rightarrow}1$ | | $1{\uparrow}2$ $1{\rightarrow}2$ |

h(m.n)

$g_l(m. n)$ $l_l(m\ n)$
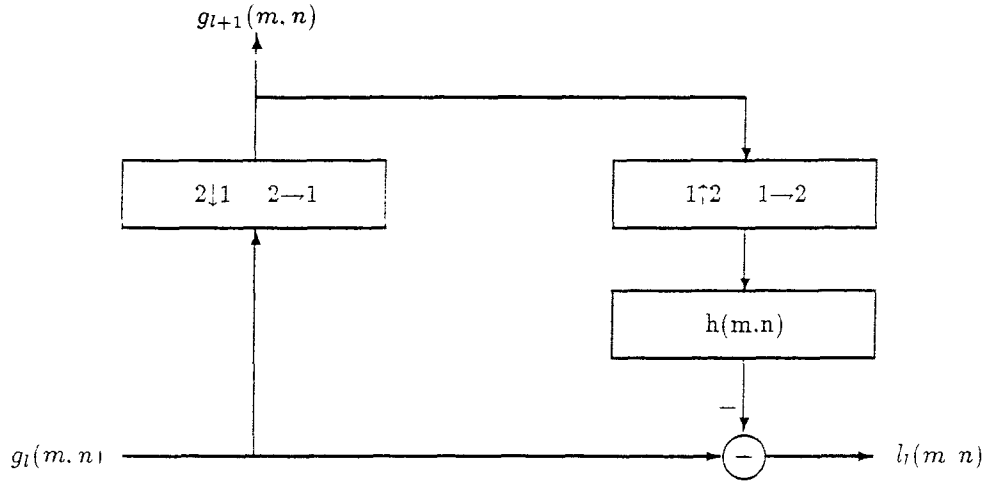
Figure 4.7:   Modification 3

4.2.4   Comparative Performances

Simulations are designed for all the filtering methods. The quantizers used are optimum quantizer. The number of pyramid levels is 4 and number of quantization levels is 3. The BPP for each pyramid are computed by the formula

$$
\begin{aligned}
BPP &= H_0 + \frac{1}{4}H_1 + \frac{1}{16}H_2 + \frac{1}{64}H_3 + \cdots + \frac{1}{2^{2(L-1)}}H_{L-1} \\
&= \sum_{l=0}^{L} \frac{1}{2^{2l}}H_l
\end{aligned}
\tag{4.12}
$$

and the SNR is the signal-to-noise ratio between the original image g_0 and the reconstructed image r_0, which is calculated as

$$
SNR = 10 \log_{10} \frac{255^2}{\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (g_0(m, n) - r_0(m. n))^2}
\tag{4.13}
$$

Table 4.1 to 4.3 list all the related simulation results for the different methods. Three images with different features are used for comparison. We have made the following observations:

52

Table 4.1: Entropies and SNR for 'woman'

Input: woman		Laplacian	Mod 1	Mod 2	Mod 3
Entropies	level 0	7.445	7.445	7.445	7.445
of	level 1	7.407	7.422	7.407	7.443
Gaussian	level 2	7.356	7.388	7.357	7.439
Images	level 3	7.280	7.315	7.280	7.407
Entropies	level 0	4.621	4.305	4.172	4.320
of	level 1	4.792	4.594	4.130	4.938
Laplacian	level 2	5.390	5.030	4.719	6.611
Images	level 3	7.280	7.315	7.280	7.407
Entropies	level 0	0.482	0.345	0.303	0.319
of	level 1	0.741	0.626	0.519	0.734
Quantized	level 2	0.951	0.814	0.652	0.998
Images	level 3	7.280	7.315	7.280	7.407
BPP of G Images		9.870	9.877	9.870	9.886
BPP of L Images		6.270	5.882	5.613	6.021
BPP of Q Images		0.841	0.666	0.587	0.680
SNR		26.187	28.310	24.121	25.747
Number of Additions		4194304	349525	2097152	2097152
Number of Multiplications		4369067	87381	2184533	2184533

- Mean pyramid is best in terms of the coding simplicity.

- Modification 2 and modification 3 are similar to each other, but modification 2 always introduces more reconstruction error.

- All three modifications reduce the number of additions and number of multiplications.

4.3   Non-Zero Interpolation

In the expansion operation shown in Figure 2.3, the Gaussian image at the higher layer is expanded in both dimensions by inserting zeros and then low-pass filtering. If the filter is designed properly, the value to be inserted does not

Table 4.2: Entropies and SNR for 'girl'

Input: girl		Laplacian	Mod 1	Mod 2	Mod 3
Entropies of Gaussian Images	level 0	7.278	7.278	7.278	7.278
	level 1	7.279	7.274	7.279	7.277
	level 2	7.282	7.267	7.281	7.270
	level 3	7.260	7.253	7.257	7.247
Entropies of Laplacian Images	level 0	3.752	3.445	3.140	3.465
	level 1	4.090	3.897	3.301	4.084
	level 2	4.844	4.516	4.057	4.800
	level 3	7.260	7.253	7.257	7.247
Entropies of Quantized Images	level 0	0.148	0.085	0.064	0.075
	level 1	0.387	0.317	0.230	0.324
	level 2	0.664	0.538	0.432	0.628
	level 3	7.260	7.253	7.257	7.247
BPP of G Images		9.666	9.664	9.666	9.665
BPP of L Images		5.190	4.815	4.333	4.899
BPP of Q Images		0.399	0.311	0.261	0.309
SNR		26.835	30.947	25.093	27.449
Number of Additions		4194304	349525	2097152	2097152
Number of Multiplications		4369067	87381	2184533	2184533

Table 4.3: Entropies and SNR for 'pepper'

Input: pepper			Laplacian	Mod 1	Mod 2	Mod 3
Entropies		level 0	6.978	6.978	6.978	6.978
of		level 1	7.128	7.111	7.128	6.981
Gaussian		level 2	7.126	7.123	7.128	6.974
Images		level 3	7.053	7.100	7.057	6.915
Entropies		level 0	5.003	4.615	4.463	4.556
of		level 1	5.499	5.269	4.896	5.470
Laplacian		level 2	5.885	5.817	5.340	6.187
Images		level 3	7.053	7.100	7.057	6.915
Entropies		level 0	0.746	0.591	0.479	0.518
of		level 1	1.079	0.991	0.780	1.081
Quantized		level 2	1.270	1.176	0.970	1.347
Images		level 3	7.053	7.100	7.057	6.915
BPP of G Images			9.315	9.312	9.315	9.267
BPP of L Images			6.856	6.131	6.021	6.418
BPP of Q Images			1.205	1.023	0.845	0.980
SNR			25.408	25.929	22.957	24.336
Number of Additions			4194304	349525	2097152	2097152
Number of Multiplications			4369067	87381	2184533	2184533

Table 4.4: BPP and SNR for 'woman' using zero and non-zero inserting

Pyramid levels	Quantizer levels	Zero-interpolation		Non-zero interpolation	
		BPP	SNR	BPP	SNR
3	3	1.250	29.663	1.184	31.006
	7	1.920	32.364	1.835	34.628
	11	2.631	34.351	2.536	38.030
	15	3.144	35.198	3.049	40.205
	19	3.203	35.434	3.104	40.845
4	3	0.970	26.514	0.899	28.614
	7	1.680	29.036	1.588	33.076
	11	2.422	30.382	2.320	36.643
	15	2.953	30.878	2.851	38.795
	19	3.026	31.026	2.919	39.640
5	3	0.904	24.475	0.831	26.830
	7	1.625	26.726	1.532	31.819
	11	2.374	27.760	2.269	35.081
	15	2.909	28.069	2.805	37.190
	19	2.983	28.170	2.875	38.081

make any difference for pyramid coding. especially when no quantizer is used. But we find this is not true when quantizers are used. The values inserted have something to do with the quality of the reconstructed image. In the following simulation we insert pixels whose gray level values are equal to left ones. The results show this interpolation can obtain about 2dB gain in SNR for the worst case. The other advantage of this scheme is that no DC offset is introduced. However, this method will slightly increase the computational complexity. Table 4.4 lists the data for image 'woman'.

Chapter 5

Conclusion

In this thesis we have analyzed two multiresolution representation approaches: Laplacian pyramid coding and pyramid subband coding using QMF. The wavelet decomposition of images, because of its current implementation is the same as subband coding with QMF, is also introduced. We found these three methods have many features in common. They all have a pyramid data structure, the images are decomposed into different resolutions, and resulting images form a coarse-to-fine image pyramid. The multiresolution representation permits the high data compression by appropriately designing the different coders for different components. The orthogonal decomposition such as subband coding, can achieve higher data compression, because they generate a sequence of uncorrelated decomposed images, or in the other words, they remove redundancy.

Subband coding and wavelet decomposition, although developed with different motivations, are implemented with similar pyramid structures and decomposition for both are obtained by QMF. Images are decomposed into orthonormal image family.

Compared with subband coding and wavelet decomposition, Laplacian pyramid coding has its advantage in implementation, because its filtering proce-

dures are simpler than the other two schemes. Laplacian pyramid also offers exact reconstruction because of its identical structure in the transmitter and receiver. The round-off in filtering computation will not result in reconstruction error as long as the filters used in the encoder and decoder are same. On the other hand. exact reconstruction of the original images through subband coding can only be obtained with infinity word length.

In Laplacian pyramid coding. the Laplacian images are generated by subtracting two Gaussian images. so the dynamic range of Laplacian image is twice the original image. The modulo limiter can be used to reduce the dynamic range without introducing any reconstruction error. The implementation of the Laplacian pyramid with modulo limiters is very easy, and does not increase the complexity of the coding scheme.

Laplacian pyramid is a versatile data structure. many related structures can be developed for different applications. The redundancy in original images can be used to simplify the pyramid structure. The 3 modifications given in Chapter 4 can reduce computational complexity greatly by eliminating one or two lowpass filters.

Laplacian pyramid coding scheme has the features similar to predictive coding system. The predict image is the reduced copy of the image at the lower layer. For generating a error image at each level (except at top of the Laplacian pyramid) the reduced image at the higher layer is first expanded and then subtracted from corresponding Gaussian image. In order to expand the reduced image in size, a row of pixels has to be inserted between two rows, and a column of pixels has to be inserted between two columns. Usually, the value inserted is zero. This value will not make too much difference if the lowpass filter is well designed. Then, we found in our simulation. a non-zero

insertion will generate a better error image and therefore reduce the quantization noise. As a result a higher SNR can be obtained. The gain from this non-zero interpolation is more significant when number of the quantization levels is large.

As we mentioned above, the proper quantization is the key to obtain data compression, and the reconstruction error is caused mainly by quantization noise. Reducing quantization noise is another consideration when using multichannel coding.

References

[1] P. J. Burt, "Multiresolution techniques for image representation. analysis. and 'smart' transmission." *Proc. SPIE Conf. on Visual Communications and Image Processing*. Vol. 1199. 1989.

[2] P. J. Burt and E. H. Adelson. "The Laplacian Pyramid as a Compact Image Code." *IEEE Trans. Commun..* Vol COM-31. No. 4, pp. 532-540.

[3] J. R. Caprio, N. Westin and J. Esposito. "Optimum Quantization for Minimum Distortion." *Proc. of the Internat'l Telemetering Conf..* pp.315-323. Nov. 1978.

[4] T. A. C. M. Classen. *et al.* "Signal Processing Method for Improving the Dynamic Range of A/D and D/A Converters," IEEE Trans. ASSP, Vol. 28, No. 5. pp.529-538. Oct. 1980.

[5] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, N.J. Prentice-Hall, 1983.

[6] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley, 1987.

[7] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Math.*, vol. 15, pp. 723-736, 1984.

[8] S. Haykin. *Communication Systems.* John Wiley and Sons, 1987.

[9] Y. Ho and A. Gersho. "A Pyramidal Image Coder with Contour-Based Interpolative Vector Quantization," *Visual Communications and Image Processing IV*, Vol. 1199. 1989.

[10] A. K. Jain. "Image Data Compression: A Review," *Preceeding of the IEEE.* Vol. 69. No. 3. Mar. 1981.

[11] N. S. Jayant and P. Noll. *Digital Coding of Waveform.* Englewood Cliffs. N.J. Prentice-Hall. 1984.

[12] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second-Generation Image-Coding Techniques." *Proceeding of the IEEE*, Vol. 73. No. 4. Apr. 1985.

[13] Z. Lin, *A Quantization Noise Feedback Technique in Pyramid Image Coding*, Master Thesis. ECE Dept. NJIT, Sept., 1990.

[14] Chung H. Lu and Jack Wang, "Quantization Noise Feedback in Pyramid Coding", p. 12.5. *1990 Picture Coding Symposium*, Cambridge, Massachusetts, USA, March 26-28, 1990.

[15] C. H. Lu. *et al*, "A DPCM System with Modulo Limiters," *Proceeding IEEE*. pp. 0581-0585. 1988.

[16] C. H. Lu, "Video Signal Enhancement Using Fine Structure Pre-Emphasis with Modulo Limiting," *Visual Communications and Image Processing*, Vol. 1199. pp.541-549. 1989.

[17] T. J. Lynch. *Data Compression Technique and Applications*. Wadsworth, 1985.

[18] B. Mahesh and W. A. Pearlman. "Image Coding on A Hexagonal Pyramid with Noise Spectrum Shaping." *Visual Communications and Image Processing IV*, Vol. 1199. pp.764-774. 1989.

[19] S. G. Mallat, "Theory for Multiresolution Signal Decomposition: The Wavelet Representation". *IEEE Trans. on Pattern Analysis and Machine Intelligence*. Vol. 11. No. 7. July 1989

[20] S. G. Mallat. "Multifrequency Channel Decomposition of Images and Wavelet Models". *IEEE Trans. on ASSP*. Vol. 37. No. 12. Dec. 1989.

[21] P. C. Millar, "Recursive Quadrature Mirror Filter: Criteria Specification and Design Method," *IEEE Trans. ASSP*. Vol. 33. pp. 413-420. Apr. 1985.

[22] K. Mizui, *et al*, "Modulo-PCM with Multi-Quantizer," Proc. ICASSP'87, pp. 1332-1335.

[23] H. G. Musmann and Peter Pirsch, "Advances in Picture Coding," *Proceeding IEEE*, Vol. 73. No. 4, Apr. 1985.

[24] A. N. Netravali and J. O. Limb, "Picture Coding: A Review," *Proceeding IEEE*, Vol 68. No 3, March 1980. pp366-407.

[25] L. O'Gorman. "A Comparison of Methods and Computation for Multi-Resolution Low- and Band-Pass Transforms for Image Processing," *Computer Vision, Graphics. and Image Processing 37.* 386-401(1987).

[26] W. K. Pratt. *Image Transmission Techniques* New York, Academic Press, 1979.

[27] V. Ramamoorthy, "A Novel Speech Coder for Medium and High Bit Rate Applications Using Modulo-PCM Principles." *IEEE on ASSP*. Vol. 33. No. 2. Apr. 1985.

[28] J. Wang, *Image Coding with Quantization Noise Feedback*, Master Thesis, ECE Dept. NJIT, May. 1989.

[29] L. Wang and M. Goldberg, "Reduced-Different pyramid: a data structure for progressive image transmission," *Optical Engineering*, Vol. 28, No. 7, July 1989.

[30] L. Wang and M. Goldberg, "Progressive Image Transmission Using Vector Quantization on Images in Pyramid Form," *IEEE Trans. on Communi-*

cations, Vol. 37, No. 12, Dec. 1989.

[31] J. Woods and S. D. O'Neil, "Subband Coding of Images," *IEEE Trans. ASSP,* Vol. ASSP-34. Oct. 1986.